# True 3D-Information from Digital Aerial Images – Edge Detection and Adjustment from an Oblique Viewing Multi-Camera System

Dennis Dahlke
Optical Information Systems
German Aerospace Center (DLR)
Berlin, Germany
Dennis.Dahlke@dlr.de

Alexander Wieden
Optical Information Systems
German Aerospace Center (DLR)
Berlin, Germany
Alexander.Wieden@dlr.de

*Abstract*— **Automated processes for building reconstruction including roof- and façade structures with accuracies better than 0.5 m have not been developed yet. This paper introduces a conceptual approach to derive 3D building edges automatically from oblique aerial imagery. A sub-pixel accurate version of the Canny algorithm is used for detection and segmentation of line segments found on seven oblique true orthomosaics (TOM) mapping a scene from different cardinal directions. With respect to the properties of oblique photogrammetric products and the characteristics of their derivation, all detected line segments are weighted based on their individual camera-target path. An adjustment approach is implemented, which adjusts redundant line segments that are found in several oblique TOM. A validation of the adjustment is done by using tacheometrically measured reference points. On average the approach yields in accuracies of 30 to 35 cm in X, Y and Z.**

## I. Introduction and Related Work

Since 2009 the Modular Airborne Camera System (MACS) is being developed and tested in the Department of Optical Information Systems at the German Aerospace Center (DLR). One product out of the MACS family is capable of capturing oblique imagery (Fig. 1), which is superior to common vertical imagery, especially when deriving façade information. Although a lot of research projects emerged in the field of recognition, reconstruction, and representation of 3D-buildings from oblique imagery over the last few years [1], [2], [3], developing photogrammetric methods for automated interpretation of oblique datasets still remain an unsolved problem. Image processing software packages like Oblivision by Idan Computers Ltd. or MultiVision by Ofek allow image analysis and simple measurements on oblique images. Predominantly those systems are made for monument documentation, façade texturing or monitoring. Contemporary solutions at the DLR go one step further. Oblique images captured by the company-owned MACS do not only serve as a basis for textures they also provide the basis for true 3D-information in terms of vectors [4]. A company-owned algorithm, the semi-global matching (SGM) [5] that has been initially designed for the creation of two-and-a-half-dimensional (2.5D) data from High Resolution Stereo Camera images are applied to the oblique data as well. Many 2.5D oblique representations of the same scene from different cardinal directions are the result, which can be seen in Fig. 2(a-g). In combination they yield in true 3D-information of a scene which has the advantage over traditional vertical photogrammetry to inherit façade information in addition to outlines of buildings respectively roof structures.

One main goal of the working group is to automate the derivation of precisely reconstructed buildings. Partial solutions have been developed in the past years. An approach for generating oblique TOM for extracting 3D geoinformation is presented in [6]. Two different and stable methods for reconstructing vertical façades from photogrammetric oblique point clouds can be found in [7] and [8]. All contribute to the process of urban 3D city modeling without a priori knowledge like cadastre maps.

The presented approach in this paper serves as a geometric basis for further interpretations and supports respectively refines the existent approaches by giving predicates about the uncertainty of each line segment and therefore its reliability. Furthermore line segments will simplify the topological linking of building façades and thus help to derive semantic information. Line segments itself serve as fundament for volume primitives.

The intention of this paper is to show the feasibility of deriving precise true 3D-line segments from several 2.5D oblique datasets.

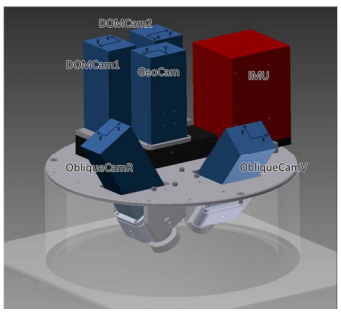


Figure 1. MACS-Oblique Camera Head.

## II. OBLIQUE DATA BASE

This section briefly introduces the concept of minimizing the dimensionality through multiple oblique TOM. 3D issues are hereby solved with several 2D datasets.

Concerning oblique imagery the 3D object space is described by several 2D projection planes. That means that metrical 2D mappings of objects are calculated by using different views, what is comparable to classical products like digital surface models or digital orthomosaics. In contrast to these traditional products a local reference (LRF) has to be defined, which is described by a cluster of oblique images. Clustered images have similar viewing geometries. In figures this means that rotation differences between images contributing to a cluster differ no more than 20°. The mean of all observations is calculated and applied as sequential rotation angles to rotation matrix $R$ (1).

$$R = \begin{pmatrix} \cos\varphi \cdot \cos\kappa & \cos\varphi \cdot \sin\kappa & \sin\varphi \\ \cos\varphi \cdot \sin\kappa + \sin\omega \cdot \sin\varphi \cdot \cos\kappa & \cos\omega \cdot \cos\kappa - \sin\omega \cdot \sin\varphi \cdot \sin\kappa & -\sin\omega \cdot \cos\varphi \\ \sin\varphi \cdot \sin\kappa - \cos\omega \cdot \sin\varphi \cdot \cos\kappa & \sin\omega \cdot \cos\kappa + \cos\omega \cdot \sin\varphi \cdot \sin\kappa & \cos\omega \cdot \cos\varphi \end{pmatrix} \quad (1)$$

In total seven clusters and rotation matrices have been created for the oblique imagery. The data products have the same properties like classical surface models and orthoimages derived from aerial or satellite imagery. Each oblique depth map is co-registered to its complementary oblique orthoimage and both are georeferenced in a LRF. Because of the rotation parameters each of the seven LRF is connected to the higher level global reference frame (GRF) and therefore also to each other. All features like points or line segments that are extracted from oblique true orthoimages in the LRF have to be consistently transformed back into the GRF. Therefore all 2D coordinates which represent features in LRF get depth information from the co registered oblique depth map. The resulting 2.5D coordinates of features are used together with the corresponding transposed rotation matrix $R^T$ to get back to the GRF (2).

$$Information(X,Y,Z)_{GRF} = Information(X,Y,Z)_{LRF} * R^T \quad (2)$$

With stringent respect to epipolarity constraints it is hereby possible to generate local datasets that can be used for further analysis like the detection of line segments. The complex problem of deriving 3D-line segments respectively their end point coordinates is therefore reduced to (2D-) image processing solutions. That is an advantage over existing methods, which mainly use epipolarity constraints to match line features from several oblique image datasets [9].

## III. GENERATION OF 3D-EDGE CANDIDATES

As has been shown in the last section it is possible to implement traditional 2D image processing solutions on several oblique TOM to detect 3D features. That is why the edge detection is performed using the Canny algorithm [10]. Canny's edge detection algorithm is restricted to single band images. All oblique TOM are color images. On one hand it is possible to generate a grayscale image by simply calculating the mean over all three bands (red, green, blue) at each pixel. Another option would be to change the color space. In this approach all TOM images are represented in HSV color space where three bands are defined by Hue, Saturation and Value. An advantage of the value-channel over common (red, green, blue) color channels is its relative independence to lighting conditions, which improves especially shadowed areas on façades in oblique TOM. In order to determine the position of an edge as precisely as possible the gradients which have been computed in the second stage of the Canny algorithm are investigated more intensively. Usually the maximum of these gradients should represent the edge pixel but to reach sub-pixel accuracies their neighborhood is taken into account. In a next step a function of second polynomial order is fitted to all perpendicular adjacent pixel values. The maximum of this function now represents the intensity maximum of an edge and therefore a more precise position. Furthermore only straight line segments are used while curved parts are discarded. Each line segment is now described by 2D coordinates for the start and endpoints in a LRF. While the edge detection step is performed on these seven oblique TOM datasets, their co-registered counterpart, the oblique surface model, will contribute the associated depth information. Line segments are now described by 2.5D coordinates for their start and endpoints in a LRF.

This procedure is repeated in all LRF. Hereby seven datasets with 2.5D line segments are generated, that can be seen in Fig. 2(h-n).

For the next step from 2.5D LRF to 3D GRF some more parameters are necessary, which are not inherent in the oblique image or surface model but given as rotation parameters for each LRF. 2.5D coordinates of all detected line segments are multiplied with the corresponding transposed rotation matrix (2) to get back into a higher level reference frame. A visualization of all detected line segments in 3D-space can be seen in Fig. 3.

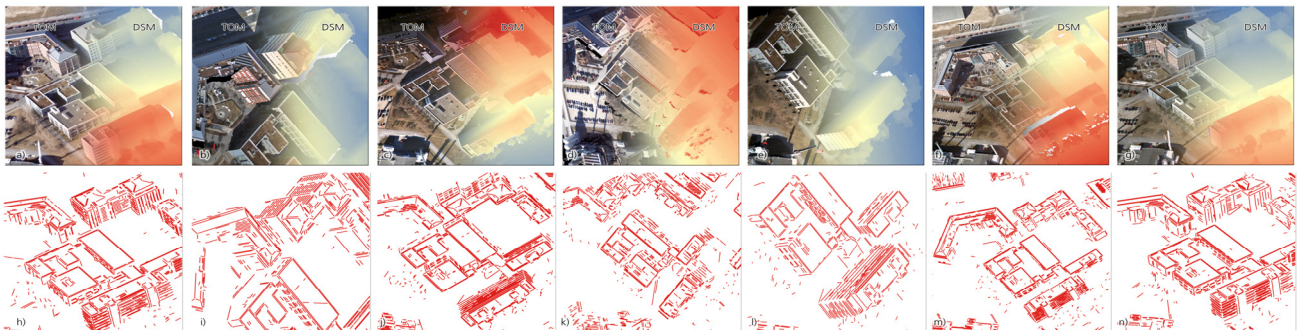


Figure 2. (a-g) Oblique true orthomosaics and surface models, (h-n) Detected line segments for seven LRF.

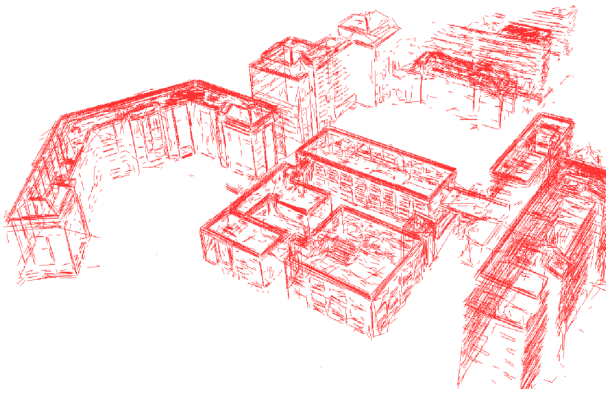

Figure 3. All detected line segments transformed back into GRF.

## IV. GROUPING AND ADJUSTMENT

An edge can be redundant if it is found in more than one LRF. In many cases the current state, with multiple combined edges, is a redundant representation of singular real world edges. Many real world edges especially those on roofs or roof constructions are found highly probable in several oblique datasets.

This section shows how redundancy of edge segments is used to for an adjustment. For example lines which are completely or partly occluded in one of the seven oblique datasets, are found in another dataset that shows this line unobstructed. Furthermore the camera-target path will be examined for each edge in order to derive weights in terms of accuracy respectively more realistic error estimates. This information will be used to minimize the variance of the weighted mean during the adjustment of redundant lines.

In a first step all line segments that describe the same real world edge are grouped. This is done by checking the line's attitude and spatial proximity. The initial criterion is the attitude of a line segment without regarding the spatial proximity. Attitude in this case means the orientation of a line segment in 3D space that is given by three angles. Based on the knowledge of each line's attitude, only lines that share a similar orientation will be checked for their spatial proximity. On one hand this prevents false matching of rectangular or somehow misaligned edges that would fulfill the spatial proximity check, and on the other hand this reduces the number of probable candidates for the computational more expensive spatial proximity check. Only line segments that differ less than 1.5° from each other remain in a group as possible candidates for a merging process.

Although all edges inside each group share a similar orientation it is very likely that they do not belong to each other because of potential great distances between them. Subsequently a spatial proximity check splits the groups into new groups with line segments that share a cylindrical buffer zone around an edge. All remaining line segments inside these separated groups have a similar orientation and their axes have a spatial proximity of no more than 50 centimeters, which empirically turns out to be a good value for a spatial proximity threshold.

In order to perform an adjustment for redundant line information, weights have to be introduced for each line segment. These weights are derived from a geometric, 3D viewpoint, namely the camera constellation.

In order to determine significant influences on the accuracy of endpoints of edges, the simplified formula (3) of the vertical case concerning the expected height accuracy [11] is adapted to the oblique case (4).

$$\sigma_{Vertical} = GSD \cdot \frac{h}{B} \cdot \sigma_{SGM} \tag{3}$$

Ground sample distance is depicted by *GSD*, height by *h*, baseline by B and $\sigma_{SGM}$ is the measuring accuracy in the SGM process. The influence of the oblique camera target path on the expected geometric accuracy of surfaces has been investigated by [7]. From there it follows that *B* is shortened in the forward looking case by the influence of the cosine of the off-nadir angle. In contrast to the vertical case *h* does not refer to height above ground but to distance between projection center and object point.

$$\sigma_{Oblique} = GSD \cdot \frac{h}{B_{Eff}} \cdot \sigma_{SGM} \tag{4}$$

In the scope of this approach the introduced weights from the camera-target path have to be decomposed in such a manner, that they are comparable and applicable for a common adjustment. Thus the oblique accuracy ($\sigma_{Oblique}$) is split up into positional ($\sigma_X$, $\sigma_Y$) and height portions ($\sigma_Z$), where $\sigma_X$, $\sigma_Y$ and $\sigma_Z$ refer to the axes of the GRF. These σ-values are calculated for each line and represent the expected accuracies of their 3D-start and -endpoint.

$$\bar{q} := \frac{\sum_{i=1}^{n}(q_i/\sigma_i^2)}{\sum_{i=1}^{n}(1/\sigma_i^2)} \tag{5}$$

The weighted mean (5) serves as input for observation equation in the adjustment, with detached 3D coordinates from all start and endpoints of a group as q and their corresponding accuracies as w. The weight matrix P comprises the accuracy information ($\sigma_X$, $\sigma_Y$ and $\sigma_Z$) for each start and endpoint.

All start and endpoints within a group are adjusted and yield in one 3D-coordinate for a start point of a line segment and one for the endpoint. Fig. 4 depicts the remaining adjusted line segments in 3D-space.

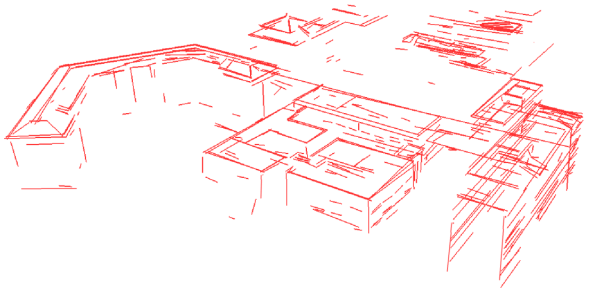

Figure 4. Remaining line segments after adjustment.

## V. RESULTS AND OUTLOOK

The adjustment result for redundant lines is verified on several chosen control points, which have been tacheometrically surveyed. All surveys in the test site have a minimum positional- respectively height accuracy of 2-5 centimeters. Ten points from this survey spread over three buildings are used to verify the outcome of the presented approach. Regarding the average accuracies the quality of the outcome is sufficient to be used for topological linking of CityGML models with a level of detail (LoD) 3, because accuracies for these edge segments can be expected to be better than 0.5 meters horizontally and vertically. The creation of LoD 2 models could be supported without reservations since all absolute accuracies are better than 1 meter. Fig. 5 shows the deviations between tacheometrically measured reference points and results of the presented approach at ten sights in the test field.

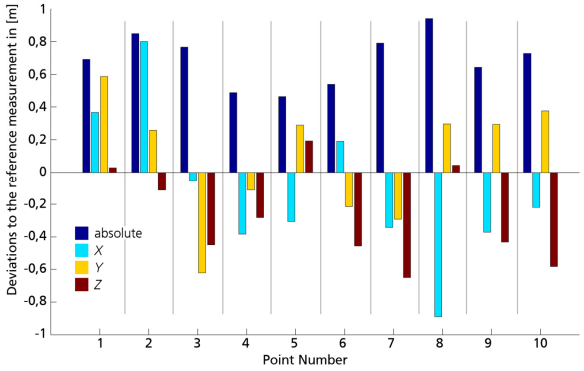


Figure 5. Deviations between reference points and result of the presented approach for ten points.

It is planned to intersect the outcome of this approach with segmented roof and façade planes, which are also derived from aerial oblique datasets.

## REFERENCES


[1] A. Alamouri, "Generation of a 3D City Model of Baalbek/Lebanon based on Historical Photos", Dissertation at the Technical University Berlin, 2011.

[2] U. Panday and M. Gerke, "Fitting of Parametric Building Models to Oblique Aerial Images", ISPRS Hannover Workshop: High – Resolution Earth Imaging for Geospatial Information , 2011.

[3] J. Xiao, M. Gerke and G. Vosselman, "Automatic Detection of Buildings with Rectangular Flat Roofs from Multi – View Oblique Imagery", ISPRS Symposium on Photogrammetry Computer Vision and Image Analysis, 2010.

[4] F. Lehmann, R. Berger, J. Brauchle, D. Hein, H. Meissner, S. Pless, B. Strackenbrock and A. Wieden, "MACS-Modular Airborne Camera System for Generating Photogrammetric High-Resolution Products", In: Photogrammetrie, Fernerkung und Geoinformation, Vol. 6, 2011, pp. 435-446

[5] H. Hirschmueller, "Semi-Global Matching – Motivation, Developments and Applications", Invited Paper at the Photogrammetric Week Stuttgart, 2011.

[6] A. Wieden and M. Linkiewicz, "True- Oblique- Orthomosaik aus Schraegluftbildern zur Extraktion von 3D-Geoinformationen", In: DGPF Tagungsband, Vol. 22, 2013, pp. in print

[7] K. Konecny, "Evaluierung der Nutzbarkeit von Punktwolken aus Schraegluftbildern zur Kartierung von Grundrisselementen", Master Thesis at the Beuth University of Apllied Sciences, Berlin, 2011.

[8] M. Linkiewicz, "Extraktion von senkrechten Fassadenebenen aus 3D-Punktwolken von Schraegluftbildern", In: DGPF Tagungsband, Vol. 22, 2013, pp. in print

[9] T. Werner and A. Zissermann, "New Techniques for Architectural Reconstruction from Photographs", In: Proceedings of the 7th European Conference on Computer Vision – Part II, 2002, pp. 541-555

[10] J. Canny, "A Computational Approach to Edge Detection", In: IEEE Transaction on Pattern Analysis and Machine Intelligence, Vol.8(6), 1986, pp. 769-798

[11] K. Kraus, "Photogrammetrie", Walter de Gruyter, Berlin, 2004.